medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Human genetic factors associated with pneumonia susceptibility, a cue for COVID-19 mortality.

Debleena Guin[1, 2#], Saroj Yadav [1, 3#], Priyanka Singh[1, 3], Pooja Singh[1, 3], Sarita Thakran[1, 3], Samiksha Kukal[1, 3], Neha Kanojia[1, 3], Priyanka Rani Paul[1, 3], Bijay Pattnaik[4], Viren Sardana[5], Sandeep Grover[6], Yasha Hasija[2], Indian Genome Variation Consortium[1], Anurag Agrawal[3,7], and Ritushree Kukreti*[1, 3]

[1]Genomics and Molecular Medicine Unit, Council of Scientific and Industrial Research (CSIR) - Institute of Genomics and Integrative Biology (IGIB), New Delhi-110007, India.

[2]Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Delhi, India.

[3]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad- 201002, India.

[4]Department of Pulmonary, Critical Care and Sleep Medicine, All India Institute of Medical Sciences, New Delhi, India.

[5] CSIR- Central Scientific Instruments Organisation, Chandigarh, India.

[6] Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen, Tübingen, Germany.

[7] Centre of Excellence for Translational Research in Asthma and Lung Diseases; Council of Scientific and Industrial Research (CSIR) - Institute of Genomics and Integrative Biology (IGIB), New Delhi-110007, India.

[#] Equal contribution to authorship

***Correspondence:**

Ritushree Kukreti
Genomics and Molecular Medicine Unit,
CSIR-Institute of Genomics and Integrative Biology (IGIB), Mall Road, Delhi-110007
E-mail: ritus@igib.res.in; ritus@igib.in
Phone no: 27662202; 011-2766 6156 Extn. 135, M – 09811689701

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

The risk for community acquired pneumonia (CAP) is partially driven by genetics. To identify the CAP-associated genetic risk loci, we performed a meta-analysis of clinically diagnosed CAP (3,310 individuals) with 2,655 healthy controls. The findings revealed *CYP1A1* variants (rs2606345, rs4646903, rs1048943) associated with pneumonia. We observed rs2606345 [G vs T; OR=1.49(1.29-1.69); p=0.0001; I^2= 15.5%], and rs1048943 [T vs G; OR= 1.31(0.90-1.71); p=0.002; I^2=19.3%] as risk markers and rs4646903 [T vs C; OR= 0.79(0.62-0.96); p=0.03; I^2=0%] as a protective marker for susceptibility to CAP, when compared with healthy controls. Our meta-analysis showed the presence of *CYP1A1* SNP alleles contributing significant risk toward pneumonia susceptibility. Interestingly, we observed a striking difference of allele frequency for rs2606345 (*CYP1A1*) among Europeans, Africans and Asians which may provide a possible link for observed variations in death due to coronavirus disease 2019 (COVID-19), a viral pneumonia. We report, for the first time, a significant positive correlation for the risk allele (T or A) of rs2606345, with a higher COVID-19 mortality rate worldwide and within a genetically heterogeneous nation like India. Mechanistically, the risk allele 'A' (rs2606345) is associated with lower expression of *CYP1A1* and presumably leads to reduced capacity for xenobiotic detoxification. We note that ambient air pollution, a powerful inducer of *CYP1A1* gene expression, is globally associated with lower, not higher mortality, as would normally be predicted. In conclusion, we find that *CYP1A1* alleles are associated with CAP mortality, presumably via altered xenobiotic metabolism. We speculate that gene-environment interactions governing *CYP1A1* expression may influence COVID-19 mortality.

Keywords: Community acquired pneumonia, *CYP1A1*, COVID-19, genetic variants, meta-analysis, air pollution.

Introduction

The ongoing coronavirus disease 2019 (COVID-19) pandemic and consequent mortality have rarely been examined through the lens of community acquired pneumonia (CAP). Pneumonia is an inflammatory condition of the lungs usually caused by bacterial or viral infection. According to the site of acquisition, pneumonia is classified as community acquired (CAP) or nosocomial (NP). The global burden of disease study (2015) stated that lower respiratory infections like pneumonia are the third most common cause of death globally [1]. Pneumonia

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

occurs when pathogens enter in the alveoli, infect, multiply and encourage a host immune response. These responses cause inflammation of the lung tissues, marking the pathogenesis of pneumonia [2]. Host genetic factors that participate in these processes starting from pathogen entry, infection, inflammation, and resolution can all be considered as good candidates in genetic association studies of pneumonia and its complications [3, 4].

A number of studies have investigated the genetic association as well as gene interaction of pneumonia risk with monooxygenase enzyme group, cytochrome P450 (CYP). In a study, a group of researcher identified CYP1A1 cytochrome P450 family 1 subfamily A member 1 (*CYP1A1*) as a critical regulator of inflammatory responses and phagocytosis in sepsis some through *CYP1A1*-involved signalling pathways that may be promising targets for treating sepsis or other inflammatory diseases [5]. Interestingly, few other studies indicated towards the role of *CYP1A1* genetic polymorphisms in infectious diseases and consequently establishing its role in inflammatory responses. Previously, it was identified that genetic variants of some host genes (*CYP1A1*, *ACE* and *IL-6*) are associated with the diversity in response to CAP [6, 7]. The selection of this gene was established based on its role in physiological and pathological processes during pneumonia infection, particularly in the immune and inflammatory responses [6]. Hence, we selected this gene. The most widely reported *CYP1A1* single nucleotide polymorphisms (SNPs) were rs2606345, rs4646903 and rs1048943. These SNPs had functional consequences which may ultimately be involved with a disease phenotype like pneumonia. The intron-located rs2606345 (C>A) determines enhanced gene expression in the presence of specific substrates (allele C) or in their absence (allele A). The minor allele of rs4646903C, located near the 3' UTR, also shows an increased inducibility. Another *CYP1A1* SNP rs1048943 (T>A, C, G) (Ile/Val), resulted in a missense amino acid substitution, is characterized by the substrate-specific increased activity for minor allele G. Thus we can suggest that genetically determined alteration of *CYP1A1* activity could contribute to lung inflammation pathogenesis. Thus, while there may be several evidence directing towards the role of this gene and/or its genetic variants in pneumonia susceptibility, so far, we lack a common mechanistic framework that integrates it all.

Here, we report a meta-analysis of pneumonia from ten included studies. We investigated the association of the *CYP1A1* gene and its genetic variants (rs2606345, rs4646903, rs1048943) with the risk of pneumonia, both CAP and NP, independently for all samples. The primary objectives of this meta-analysis are: (1) to determine a precise numerical estimate of the

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

impact of *CYP1A1* risk allele and pneumonia risk, (2) to determine whether age or sex alters the odds of incident pneumonia in presence of the *CYP1A1* SNP alleles (3) to determine whether any association between *CYP1A1* and CAP is generalizable to COVID-19. We hypothesized that *CYP1A1* genetic variant would increase the odds of incident pneumonia and COVID-19 mortality.

Results

The current systematic search included eleven studies totalling 5,965 unique subjects (3,310 cases and 2655 healthy controls). The flow chart for the study selection process is represented in **Figure 1**. Demographic characteristics and clinical details of all the included studies are provided in **Supplementary table 1**. The study populations primarily comprised of well-characterized cohorts in Russia [7-15] and one included Chinese cohort [6]. The mean age of all the pooled participants was 29.33 years (31.16 years for cases and 27.49 years for controls). All the studies discuss the association of *CYP1A1* genetic variants (rs2606345, rs4646903, rs1048943) with the risk of pneumonia. For cumulative quality assessment, three of ten articles were deemed as good quality (cut off score of ≥7), six articles (≥5–6 score) were categorised under moderate, and finally any scores below 5 were judged as poor quality which included one article **(Supplementary table 2)**.

This meta-analysis has been categorised into two types of pneumonia, CAP and NP, based on the cases included in each study (**Table 1**). Among all the ten studies included comparing patients with pneumonia with healthy controls, ten provided the data on rs2606345, ten provided data on rs1048943 and eight provided the data on rs4646903. Of all the *CYP1A1* variants studied, we observed the most significant distribution of rs2606345, rs4646903, and rs1048943, with data available for 1,336, 1,382, and 909 individuals with CAP and 632, 508, and 449 individuals with NP, respectively. Overall, statistically significant heterogeneity ($P<0.10$, $I^2<50\%$) was observed among these included studies, therefore the fixed effect models were used for the following analysis. Our meta-analysis demonstrated that *CYP1A1* genetic polymorphisms significantly correlated with the increased risk of pneumonia under the allele model rs2606345 [G vs T; OR=1.49(1.29-1.69); p=0.0001; I^2= 15.5%], and rs1048943 [T vs G; OR= 1.31(0.90-1.71); p=0.002; I^2=19.3%] as risk markers and rs4646903 [T vs C; OR= 0.79(0.62-0.96); p=0.03; I^2=0%] as a protective marker for susceptibility to CAP, when compared with healthy controls.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Community acquired pneumonia (CAP)

The pooled OR derived from 3,310 cases and 2,655 healthy controls subjects in 10 recruited studies was statistically significant (see **Table 1**). Homogeneity analysis for the ORs from the ten studies of the risk allele suggested that there was statistically significant evidence for heterogeneity of the odds ratio (ORs) among the other groups. The genetic variant, rs2606345 and rs1048943 were significantly overrepresented in CAP patients with ORs of 1.49(1.29-1.69); p=0.0001, and 1.31(0.90-1.71); p=0.002; respectively and qualified as risk markers (**Figure 2 A** and **C**). In contrast, rs4646903 was observed to be protective (0.79(0.62-0.96); p=0.03) (**Figure 2 B**).

Nosocomial pneumonia (NP)

The pooled OR derived from 3,310 cases and 2,655 healthy controls subjects in ten recruited studies was not statistically significant (**Supplementary table 1**). The risk allele of all the three *CYP1A1* genetic variants were seen to have increased frequency in NP patients as compared to healthy controls, however no statistically significant correlation was identified between the polymorphisms and the risk of NP. .

Test for publication bias

We failed to observe publication bias in any of the associations reported above, showing that our results were reliable (**Table 1 and Supplementary figure 1**).

Genetic variability of rs2606345 in global populations and a possible link to COVID-19

Since rs2606345 is the major allele in Europeans (66.6%) but not in other populations (African 5%, Asian 5-30%) [16], we explored a possible association with high variability in regional mortality of the ongoing COVID-19 pandemic, through a spatial analysis of the association between COVID-19 mortality with the alternate allele of rs2606345 (*CYP1A1*). To estimate the correlation pattern we performed a linear model regression and Pearson's correlation coefficient exploring the variation of COVID-19 mortality with the allele frequency data (for rs2606345) for each country (**Figure 3, A** and **B**). The regression analysis showed a strong positive correlation ($p < 0.22 \times 10^{-16}$) of higher mortality with higher frequency of A (or T), with a large effect size goodness of fit (r^2) ~ 43.5% (**Figure 3 D**). The spatial analysis in the genetically heterogeneous Indian population also confirmed a similar trend (**Figure 4 A** and **B**), with regions that have European genetic ancestry being more

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

affected [17, 18]. The linear model analysis revealed a positive correlation between mortality related to COVID-19 and alternate allele frequency ($p < 0.039$, $r^2 = 18.8\%$) (**Figure 4 C**).

Additionally, air pollution is known to upregulate *CYP1A1* expression [19], similar to what the protective allele is expected to do, we estimated the correlation between the population-weighed average level of exposure to concentration of suspended particulate matter in air measuring less than 2.5 microns in diameter ($PM_{2.5}$) (**Figure 3 C**) with COVID-19 mortality. We observed an inverse relationship (r=-0.35) between the two. **Figure 3 E** represents the significant negative correlation ($p < 2.15 \times 10^{-6}$) with the goodness of fit (r^2) estimated to be 12.3%. This suggests additional gene-environment interactions. In view of the preliminary nature of the evidence, this was not explored further at this point.

Discussion

This study is an attempt to investigate the role of host genetic factors in pneumonia susceptibility. Like pneumonia, most infectious diseases, are a consequence of a complex network of host genetics and pathogen genetic factors that may be inducible by several non-genetic factors. The host immune mediated response determines the course of the disease, its susceptibility, progression, and severity. Such immune response related genes may serve as good candidates in establishing a genetic association in infectious diseases. Therefore, we considered xenobiotic detoxification genes encoding, CYP enzymes, which are also involved in inflammatory responses by inducing oxidative stress on encountering foreign bodies[20]. This meta-analysis reiterates prior observations that suggest a genetic contribution of *CYP1A1* to pneumonia risk, and provides a more precise estimate of the risk of individuals carrying *CYP1A1* genetic variants (rs2606345, rs4646903, rs1048943) for developing pneumonia.

Our study revealed that the alternate allele (A in plus-strand or T in minus-strand) of rs2606345 increased pneumonia susceptibility in Russian population. We also observed, this allele to be the major allele in European (66.6%) and Russian (~80%) population unlike in the other populations (African 5%, Asian 5-30%, American 39%) [16]. Interestingly, we observed a striking similarity in the trend for the recent outbreak of COVID-19, a global pandemic. We would also like to warn that since this is an ongoing pandemic, the numbers are changing with time and this is circumstantial evidence. With the second surge

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

progressing, we still observe similar trends with higher COVID-19 mortality in populations with higher A allele frequency (rs2606345) (**Figure 3 B** and **D**) Likewise, the spatial analysis in the Indian population also directed towards a similar trend, where the frequency of alternate allele of this SNP (rs2606345) varied from 9% (in sub-population of northern Madhya Pradesh) to 60% (in sub-population of Maharashtra) [17] (**Figure 4 A** and **B**). The highest number of COVID-19 deaths have been observed in the west-Indian state of Maharashtra, where allele frequency of A ranges between 30-60% [17]. However, this is an evolving situation, as India goes through a ferocious second wave. It is however noted that with a median risk allele frequency of over 30%, India is an at-risk nation, with only some regions expected to have genetic protection.

The COVID-19, a viral pneumonia [21, 22], is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [23-26]. The disease is caused by host inflammatory responses affecting the lungs and lung blood vessels [27]. Physiologically, the virus enter into the host system by attaching to angiotensin converting enzyme 2 (ACE2) receptors [28]. The viral RNA, on release, hijacks the host cell's machinery to initiate viral replication and spread. The SARS-CoV-2 infection, triggers both innate and adaptive immune responses (both humoral and cell-mediated immunity). In the later stage when the infection spreads the epithelial-endothelial barrier integrity of lungs is compromised. It also infects the lung capillary endothelial cells, producing a rush of inflammatory response by influx of an army of immune cells like monocytes, neutrophils, macrophages infiltrating to the alveolar spaces [29]. This activates the host immune system to release the cytokines, several inflammatory cytokines (IL-1, IL-6, TNFα, and IFNγ) [30]. Interestingly, the CYP P450 enzyme activity has been involved in the inflammatory response [31, 32]. Fang *et al*. first observed a significant reduction in *CYP1A1* expression in pigs infected with *Mycoplasma hyopneumoniae* compared with the naive controls lung [33]. Later, a report validated that CYP1A1 suppress the inflammatory response caused by *M. hyopneumoniae* infection, via PPAR-signaling pathway [32]. Tian Li X *et al*. demonstrated on bacterial lipopolysaccharide (LPS) stimulation, *CYP1A1* gene was upregulated, resulting in increased production of pro-inflammatory cytokines like TNF-α and IL-6 [5]. They also suggested that this upregulation reduced the phagocytosis of the bacteria in the macrophages by decreasing the expression of SR-A, a macrophage channel crucial for phagocytosis. As a target regulated by *CYP1A1*, PPAR-γ inhibits the expression of inflammatory activating factors along with NF-κB factors [32] and increased *CYP1A1* expression is repressed by inhibition of NF-κB [34]. This indicated that

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

CYP1A1 expression in pulmonary macrophage is vital for host defence via regulating macrophages phagocytosis. Meanwhile, the production of pro-inflammatory cytokines drastically downregulate *CYP1A1*, disrupting the balance between the CYP enzymes, producing reactive oxygen species (ROS) thereby increasing the oxidative stress and triggering apoptosis [35, 36]. These pro-inflammatory cytokines inhibit apoptotic cell clearance in the lung, worsening the inflammation [37, 38]. Although the exact physiological role of *CYP1A1* with respect to SAR-CoV-2 has not been established, this genetic association with the current pandemic is plausible based on available data from other systems.

We note that there can be several other non-genetic factors that may increase the disease incidence. Demographic and socio-economic factors may account for COVID-19 case fatality rate (CFR) variations [39, 40]. Here, we focussed on environmental factors that may be indirectly interacting with *CYP1A1*. While there have been concerns about air pollution leading to increased COVID-19 deaths [41], the most affected nations seem to be those with the best air quality [40]. In one of our previous findings, we have shown that the A allele (rs2606345) is responsible for a 70-80% reduction in promoter activity, which would reduce enzyme activity [17]. As a caveat, induction of *CYP1A1* would be expected to be protective. This interpretation could explain counterintuitive lower mortality in regions with higher ambient air pollution, since pollutants are known to be a powerful inducer of *CYP1A1* gene expression [42]. The unexpected associations of smoking (induces *CYP1A1* gene expression [43, 44] with COVID-19 severity [45] and death [46] could also be understood in this framework. It is clearly noted that both air pollution and smoking are extremely detrimental to health and overall increase respiratory as well as non-respiratory morbidity and mortality.

To conclude, we find that *CYP1A1* alleles are associated with CAP mortality, presumably via altered xenobiotic metabolism. We speculate that gene-environment interactions governing *CYP1A1* expression may influence COVID-19 mortality. Towards this by-product of the main conclusion, we fully acknowledge there are several other factors like demographics (average age of population, population density, sex of a person, ethnic diversity, genetic variability), comorbidities (cardiovascular, cancer, and chronic respiratory diseases), socio-economic factors (GDP per capita, healthcare infrastructure), and political regime (Govt. isolation policies, social distancing, stringency index) that contribute to SARS-CoV-2 mortality and could be potential confounders. Further, the uncertainty of estimation of

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

mortality, while the pandemic is still on, the limitations of the meta-analysis in itself, with limited population data and sample size are some major pitfalls.

Materials and Methods

This meta-analysis was conducted as per the recommendations of the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines [47]. The selected studies were those in which the relationship between *CYP1A1* gene polymorphisms and risk of pneumonia disease has been evaluated. Bibliographic databases like MEDLINE (PubMed), Web of Science, and Cochrane database of systematic reviews were searched for all articles published till January 13, 2020. The keywords used to identify relevant studies were "*CYP1A1*", "Pneumonia", "genetic variant", and "single nucleotide polymorphisms" using AND/ OR Boolean operators. Cross references of each study retrieved were also examined for inclusion in case they discuss the effect of *CYP1A1* genetic variant and its risk in pneumonia.

Two investigators independently reviewed each study for its inclusion in the meta-analysis. The inclusion criteria were: (1) studies conducted on human population only, (2) included the effect of *CYP1A1* genetic variant with available numeric data, (3) adopted a case-control study design (4) provided a detailed assay method. Studies (1) any other type of lung inflammation apart from pneumonia or pneumonia as a consequence of any exposure or pneumonia existing with comorbid conditions, (2) no defined diagnostic criteria for pneumonia, and (3) genotypic data not in accordance with Hardy Weinberg equilibrium were excluded.

Allele frequency data extracted for each case and control were extracted into contingency tables to calculate the odds of pneumonia in patients carrying the risk allele of the associated variants. In case of missing allele frequency data, the corresponding OR and p value were calculated from genotypic data given. The references of the retrieved articles were manually screened to identify additional studies. In case of studies where genotypic data is given allelic data is calculated to maintain a consensus across studies. The included studies and their characteristics like first author, year of publication, population, disease, genetic variant, odds ratio, genotyping method, risk allele, sample size (cases and control) and male female distribution and their quality assessment score. All the articles described some variant of

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

pneumonia infection, one study discussing *Mycoplasma pneumoniae* infection [6], community acquired pneumonia (CAP) [8, 11, 12, 15, 48], nosocomial pneumonia (NP) [7, 49], both CAP and NP [9, 14] and relapsing pneumonia [10]. Since *M. pneumoniae* infection is the most common form of CAP, the data from this article is included of CAP cohort. Patients having frequently recurring (relapsing) pneumonia (J18, according to the ICD-10) with unspecific organism of infection, this has also been included for CAP for easy interpretation. The included cases were diagnosed by experienced professionals based on symptoms, medical histories and the clinical, radiology or laboratory results (chest X-ray, spirometry measures, etc.), and physical examination of new lung infiltrates or lower respiratory tract infection. Controls were age and gender matched healthy volunteers with no previous history of relevant infectious diseases.

Stata 16.0 (Stata Corporation, College Station, TX) [50, 51] was used to generate pooled ORs using inverse variance-weighted, fixed effects meta-analysis [52, 53]. Heterogeneity of data was evaluated using the I^2 statistic, with I^2 greater than 50% considered significant heterogeneity [54]. Summary ORs were represented as a point estimate and 95% confidence intervals (CIs) on a forest plot [55], and publication bias was evaluated [56-58]. The methodological assessment of all the selected articles was performed by two reviewers independently using the modified Newcastle-Ottawa Scale (NOS) for non-randomised studies [59]. The quality weight was assigned on the basis of eight categories primarily based on three broad criteria: selection of study groups; comparability of the groups; and ascertainment of either the exposure or outcome of interest for case-control, respectively. A maximum of one star was awarded for each detail present in the study for each parameter except for comparability, where a maximum of two stars can be given. A cumulative score of the number of stars obtained for each study reflected its quality. In case of conflicting scores, a consensus was reached upon discussing with another author. A study was regarded as a high-quality study when it rated six or more stars.

We performed a spatial analysis (among countries) to find the association between COVID-19 mortality with the alternate allele of rs2606345 (*CYP1A1*). We used the online resource https://ourworldindata.org/ to retrieve data on COVID-19 mortality (death per million) for each country as on May 24, 2021. We also included air pollution data obtained as a level of exposure to concentration of suspended particulate matter in air measuring less than 2.5 microns in diameter ($PM_{2.5}$) retrieved on till May 24, 2021. The population specific allele

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

frequency data for rs2606345 was obtained from 1000genome browser [16] on March 08, 2021. For the Indian population data, the recorded COVID-19 related death was retrieved from the ministry of health and family welfare (https://www.mohfw.gov.in/) as on 24 May 2021. The regional case fatality rate (CFR) was calculated as total number of COVID-19 related death upon a sum of total death and total patients recovered [60]. The R version 3.4.2 (R Project for Statistical Computing) was used for statistical analyses and spatial analysis plots. Initially, the data cleaning process was performed wherein we remove the missing values from the two datasets. The variable of interest includes total death per millions and SNP (rs2606345) 'A' allele frequency. We first estimated if COVID-19 mortality among countries varied with the allele frequency. For this, we conducted a linear regression and calculated the Pearson's correlation coefficient. Pearson's correlation analysis to investigate the linear or nonlinear relationship between two continuous variables. Utilizing this association, a simple linear regression model was applied to measure the relationship between mortality (dependent variable) and 'A' allele frequency of rs2606345 (independent variable). The goodness of fit (r^2) was evaluated to determine the model ($p < 0.05$).

SUPPLEMENTARY MATERIALS

The Supplementary Material for this article can be found online at

AUTHOR CONTRIBUTIONS

RK devised the concept of the review, reviewed the manuscript and supervised the overall study till final manuscript preparation. DG and SY performed the literature search independently and, reviewed articles and, extracted the data. DG generated the forest and funnel plots. PS (Priyanka Singh) performed the linear regression and calculated Pearson's correlation and prepared all the spatial analysis images. SG helped us through the meta-analysis and running the statistical analysis through STATA. DG and PS (Pooja Singh) performed the quality assessment of the included articles and prepared the supplementary material. Upon discrepancies, RK resolved the conflict. DG and SY wrote the manuscript. ST cross-checked all the data, re-calculated the numerical values and statistical analysis. NK, SK, PRP helped in manuscript writing, reference management, and preparation of tables. RK and AA reviewed the manuscript, figures and tables. DG, SY and NK edited the manuscript. SY and NK prepared the tables. RK, YH, and BP finally revised the figures, tables and the whole manuscript. VS improved the writing of the overall manuscript. RK supervised the whole study. AA gave valuable inputs regarding the COVID-19 mortality rate and directed us

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

to estimate the inter-ethnic difference for COVID-19 mortality. AA reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

Financial support for this research work has been provided by the Council of Scientific and Industrial Research (CSIR), grant number OLP1154.

ACKNOWLEDGEMENTS

We are highly indebted to our founder director, Prof. S.K. Brahmachari, whose vision in the field of genomics paved us a path. SY, PS (Pooja Singh), ST, NK, and PRP acknowledge CSIR, Government of India for providing their fellowships. DG, PS (Priyanka Singh) and SK acknowledges ICMR, UGC and DBT, Govt. of India, for their financial assistance, respectively. We thank the anonymous reviewers for their helpful suggestions in improving the manuscript. The authors are thankful to Dr L.E. Salnikova, V. A. Negovsky Research Institute of General Reanimatology, Russian Academy of Medical Sciences, Moscow Neurology Division, for sharing genotypic and allelic data from their patient population.

CONFLICTS OF INTERESTS

The authors declare no conflict of interest.

LIST OF FIGURES AND TABLES

1. **Figure 1:** Flow chart of study selection in metaanalysis of CYP1A1 polymorphisms with Pneumonia risk.Study methodology for the inclusion and exclusion of studies exploring the role of *CYP1A1* genetic variants in pneumonia patients. The number of studies excluded on each step is represented as N.

2. **Figure 2:** Forest plot determining pneumonia risk associated of *CYP1A1* genetic variants: a) for rs2606345, b) rs4646903 c)1048943; in community acquired pneumonia (CAP). The square and horizontal lines correspond to the study- specific odds ratio(OR) and 95% confidence interval (CI). The area of the square refers to the study specific weight (inverse of variance).The diamond represents the summary of OR and 95% CI.

3. **Figure 3:** A contrasting feature between allele frequency of rs2606345 (*CYP1A1*), COVID-19 mortality rate, and particulate matter air pollution worldwide. **A.** Geospatial frequency maps depicting the distribution of minor allele frequency of rs2606345 (*CYP1A1*) SNPs. **B.** A geospatial density map of the number of deaths per million inhabitants as on 24 May 2021 worldwide due to COVID-19. **C.** A geospatial density map of the number of air quality index (AQI) $PM_{2.5}$ worldwide obtained from ourworldindata.org till 2016. **D.** The regression analyses showing the goodness of fit and Pearson correlation coefficient for the allele frequency and COVID-19 death per million. Some of the values have been jittered by drawing samples from uniform distribution. **E**. The regression analyses showing the goodness of fit and Pearson

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

correlation coefficient for the AQI $PM_{2.5}$ and COVID-19 case fatality ratio across Indian states. COVID-19 mortality is extracted from ourworldindata.org till 24 May 2021. The minor allele frequency distribution is plotted for SNPs across global populations of rs2606345. The data is obtained from population frequency data of the 1000genome browser on 8 March, 2021. Air quality is represented by concentration of particulate matter of size 2.5 micrometres (µm) ($PM_{2.5}$). The data is obtained from ourworldindata.org till 2016. The white coloured areas in the map show the absence of data. A half open intervals includes only one of its end-points and is denoted by mixing notations for open and closed intervals. For e.g., (0-1] means greater than 0 and less than or equal to 1 and [0,1) means greater than or equal to 0 and less than 1.

4. **Figure 4:** A contrasting feature between COVID-19 mortality rate and allele frequency data of rs2606345 (*CYP1A1*) in Indian population. **A.** Indian population state-wise frequency of case fatality rate (CFR) for COVID-19. CFR data given till 24 May, 2021. **B.** Allelic 'A' distribution of rs2606345 variant in major ethnic groups in the Indian population. **C.** The regression analyses showing the goodness of fit and Pearson correlation coefficient for the allele frequency and COVID-19 case fatality ratio across Indian states. The COVID-19 case fatality ratio has been extracted from MoHFW.gov.in. Here, CFR is calculated as number of death upon number of death + number of recovered from COVID-19 infection, since COVID-19 is an ongoing pandemic [60]. A half open intervals includes only one of its end-points and is denoted by mixing notations for open and closed intervals. For e.g., (0-1] means greater than 0 and less than or equal to 1 and [0,1) means greater than or equal to 0 and less than 1. Figure 4b reproduced with permission [17].

5. **Table 1:** Pooled odds ratio for allelic comparisons for studies exploring association of *CYP1A1* variants- rs2606345, rs4646903, rs1048943 in patients with risk of CAP.

Table 1: Pooled odds ratio for allelic comparisons for studies exploring association of *CYP1A1* variants- rs2606345, rs4646903, rs1048943 in patients with risk of CAP.

Gene (SNP)	Risk allele	No of studies	Population	Total samples	CAP patients		Total	Control		Total	OR(95%CI)	p value	I^2	Model	Test of publication bias
					Risk allele present	Risk allele absent		Risk allele present	Risk allele absent						Egger's test
CYP1A1 (rs2606345)	T	5	Chinese, Russian	2691	2112	560	2672	1876	834	2710	1.49(1.29-1.69)	**0.0001**	18.88	F	0.361
CYP1A1 (rs4646903)	C	3	Russian	1929	168	1650	1818	231	1809	2040	0.79(0.62-0.96)	**0.03**	0.00	F	0.948
CYP1A1 (rs1048943)	G	6	Russian	2469	111	2007	2118	99	2721	2820	1.31(0.90-1.71)	**0.002**	0.00	F	0.002

Bold characters highlight significantly associated alleles with respective P values. CAP, community acquired pneumonia; OR, odds ratio; CI, confidence interval; F, Fixed effect model. All p values calculated using chi-square test

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

REFERENCES

1. WHO Factsheet, The top 10 causes of death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

2. Jain V, Vashisht R, Yilmaz G, et al. Pneumonia Pathology. [Updated 2021 Apr 1]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK526116/

3. Kumar V, Wijmenga C, and Xavier RJ. Genetics of immune-mediated disorders: from genome-wide association to molecular mechanism. Curr Opin Immunol. 2014, 31, 51-57.

4. Cooke G S, Hill A V, Genetics of susceptibility to human infectious disease. Nat. Rev. Genet. 2001, 2, 967-977.

5. Tian LX, Tang X, Zhu JY, Luo L, Ma XY, Cheng SW, Zhang W, Tang WQ, Ma W, Yang X, Lv CZ, Liang HP. Cytochrome P450 1A1 enhances inflammatory responses and impedes phagocytosis of bacteria in macrophages during sepsis. Cell Commun Signal. 2020, 18, (1), 70.

6. Zhao J, Zhang W, Shen L, Yang X, Liu Y, Gai Z. Association of the ACE, GSTM1, IL-6, NOS3, and CYP1A1 polymorphisms with susceptibility of mycoplasma pneumoniae pneumonia in Chinese children. Medicine 2017, 96, (15), e6642.

7. Salnikova LE, Smelaya T V, Vesnina I N, Golubev A M, Moroz V V. Genetic susceptibility to nosocomial pneumonia, acute respiratory distress syndrome and poor outcome in patients at risk of critical illness. Inflammation 2014, 37, (2), 295-305.

8. Salnikova LE, Smelaya TV, Moroz V V, Golubev A M, Rubanovich A V, Host genetic risk factors for community-acquired pneumonia. Gene 2013, 518, (2), 449-56.

9. Salnikova LE, Bold TV, Moroz VV, Golubev AM, Panasenkov NH, Khomenko RV, Kharlamov IV, Lapteva NS, Kuznetsov GI, Poroshenko G, Rubanovich AV, Xenobiotic Detoxification Genes and Their Role in the Development of Pneumonia. OboireAnimatologi 2008, 4, 6.

10. Korytina, GF, Yanbaeva DG, Babenkova LI, Etkina EI, Victorova TV. Genetic polymorphisms in the cytochromes P-450 (1A1, 2E1), microsomal epoxide hydrolase and glutathione S-transferase M1, T1, and P1 genes, and their relationship with chronic bronchitis and relapsing pneumonia in children. Journal of molecular medicine (Berlin, Germany) 2005, 83, (9), 700-10.

11. Salnikova LE, Bold TV, Moroz VV, Golubev M, Lapteva NS, Poroshenko G, Rubanovich AV, Genetic predisposition to the development of acute community acquired pneumonia. OboireAnimatologi 2010, 6, 1.

12. Bold TV, Salnikova LE, Moroz VV, Golubev AM, Zarzecki V, Rubanovich AV, Genetic polymorphism and the incidence of complication of pneumonia of various genesis. OboireAnimatologi 2011, 7, 2.

13. Moroz V V, Smelaya TV, Salnikova L.E, Golubev AM, Rubanovich AV. Study of the effect of functional polymorphism of xenobiotic detoxification genes and the renin-angiotensin system on the development of acute community-acquired pneumonia. Vestn Ross Akad Med Nauk. 2011;(11):12-6. Russian.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

14. Salnikova LE, Smelaya TV, Moroz VV, Golubev AM, Rubanovich AV. Functional polymorphisms in the CYP1A1, ACE, and IL-6 genes contribute to susceptibility to community-acquired and nosocomial pneumonia. Int J Infect Dis 2013, 17, (6), e433-42.

15. Salnikova LE, Smelaya TV, Golubev AM, Rubanovich AV, Moroz VV. CYP1A1, GCLC, AGT, AGTR1 gene-gene interactions in community-acquired pneumonia pulmonary complications. Molecular biology reports 2013, 40, (11), 6163-76.

16. 1000 Genomes Project Consortium, Auton A, Brooks LD et al. A global reference for human genetic variation. Nature 2015, 526, 68-74.

17. Talwar P, Kanojia N, Mahendru S, et al. Genetic contribution of CYP1A1 variant on treatment outcome in epilepsy patients: a functional and interethnic perspective. Pharmacogenomics J. 2017, 17, (3), 242-251.

18. Jain A, Bhoyar RC, Pandhare K, et al. IndiGenomes: a comprehensive resource of genetic variants from over 1000 Indian genomes. Nucleic Acids Res. 2021, 49, (D1), D1225-D1232.

19. Totlandsdal AI, C. F., Schwarze P, Refsnes M, Låg M. Diesel exhaust particles induce CYP1A1 and pro-inflammatory responses via differential pathways in human bronchial epithelial cells. Part Fibre Toxicol. 2010, 7, 41.

20. Stading R, Chu C, Couroucli X, Lingappan K, Moorthy B. et al. Molecular role of cytochrome P4501A enzymes in oxidative stress. Curr Opin Toxicol. 2020, 20-21, 77-84.

21. Berlin D A, Gulick RM, and Martinez FJ. Severe Covid-19. N. Engl. J. Med. 2020, 383, 2451-2460.

22. Gandhi RT, Lynch JB and Del Rio C. Mild or Moderate Covid-19. N. Engl. J. Med. 2020, 383, 1757-1766.

23. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, and Chang H. The Architecture of SARS-CoV-2 Transcriptome. Cell 2020, 181, (4), 914-921.e10.

24. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, et al. A new coronavirus associated with human respiratory disease in China. Nature 2020, 579, (7798), 265-269.

25. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, Zhang J, Weng T, Zhang Z, Wu Z, et al. Molecular Architecture of the SARS-CoV-2 Virus. Cell 2020, 183, (3), 730-738.e13.

26. Zhang YZ, and Holmes EC. A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. Cell 2020, 181, (2), 223-227.

27. Dorward DA, Russell CD, Um IH, et al.Tissue-Specific Immunopathology in Fatal COVID-19. Am J Respir Crit Care Med. 2021 Jan 15;203(2):192-201.

28. Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, Wang Q, Xu Y, Li M, Li X, Zheng M, Chen L, Li H. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharm Sin B. 2020 May;10(5):766-788.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

29. Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, Liu S, Zhao P, Liu H, Zhu L, Tai Y, Bai C, Gao T, Song J, Xia P, Dong J, Zhao J, Wang FS. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. Lancet Respir Med. 2020 Apr;8(4):420-422.

30. Li J, Fan JG. Characteristics and Mechanism of Liver Injury in 2019 Coronavirus Disease. J Clin Transl Hepatol. 2020 Mar 28;8(1):13-17.

31. Projean D, Dautrey S, Vu HK, Groblewski T, Brazier JL, Ducharme J. Selective downregulation of hepatic cytochrome P450 expression and activity in a rat model of inflammatory pain. Pharm Res. 2005 Jan;22(1):62-70.

32. Fang X, Zhao W, Xu J, Tu F, Wang X, Li B, Fu Y, Ren S. CYP1A1 mediates the suppression of major inflammatory cytokines in pulmonary alveolar macrophage (PAM) cell lines caused by Mycoplasma hyponeumoniae. Dev Comp Immunol. 2016 Dec;65:132-138.

33. Fang X M, Zhao WM, Fu YF, Tu F, Li BX, Wang XM, Zhao F, Ren SW. Difference in susceptibility to mycoplasma pneumonia among various pig breeds and its molecular genetic basis. Scientia Agricultura Sinica 2015, 48, (14), 2839-2847.

34. Vogel CF, Khan EM, Leung PS, Gershwin ME, Chang WL, Wu D, Haarmann-Stemmann T, Hoffmann A, Denison MS. Cross-talk between aryl hydrocarbon receptor and the inflammatory response: a role for nuclear factor-κB. J Biol Chem. 2014 Jan 17;289(3):1866-75.

35. Bui P, Imaizumi S, Beedanagari SR, Reddy ST, Hankinson O. Human CYP2S1 metabolizes cyclooxygenase- and lipoxygenase-derived eicosanoids. Drug Metab Dispos. 2011 Feb;39(2):180-90.

36. Umannová L, Machala M, Topinka J, Schmuczerová J, Krčmář P, Neča J, Šujanová K, Kozubík A, Vondráček J. Benzo[a]pyrene and tumor necrosis factor-α coordinately increase genotoxic damage and the production of proinflammatory mediators in alveolar epithelial type II cells. Toxicol Lett. 2011 Oct 10;206(2):121-9.

37. Borges VM, V. R., McPhillips KA, Kench JA, Morimoto K, Groshong SD, Richens TR, Graham BB, Muldrow AM, Van Heule L, Henson PM, Janssen WJ TNFa inhibits apoptotic cell clearance in the lung, exacerbating acute inflammation. Am J Physiol Lung Cell Mol Physiol 2009, 297, (4), L586-L595.

38. Schmidt EP, Tuder RM. Role of Apoptosis in Amplifying Inflammatory Responses in Lung Diseases. J Cell Death. 2010 Jul 22;3:41-53.

39. Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. Proc Natl Acad Sci U S A. 2020 Sep 8;117(36):22035-22041.

40. Sorci G, Faivre B, Morand S. Explaining among-country variation in COVID-19 case fatality rate. Sci Rep. 2020 Nov 3;10(1):18909.

41. Venter ZS, A. K., Chowdhury S, Lelieveld J. , COVID-19 lockdowns cause global air pollution declines. . Proc Natl Acad Sci U S A. 2020, 117, (32), 18984-18990.

42. Ibrahim M, MacFarlane EM, Matteo G, Hoyeck MP, Rick KRC, Farokhi S, Copley CM, O'Dwyer S, Bruin JE. Functional cytochrome P450 1A enzymes are induced in mouse and human islets following pollutant exposure. Diabetologia. 2020 Jan;63(1):162-178.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

43. Hussain T, Al-Attas OS, Al-Daghri NM, Mohammed AA, De Rosas E, Ibrahim S, Vinodson B, Ansari MG, El-Din KI. Induction of CYP1A1, CYP1A2, CYP1B1, increased oxidative stress and inflammation in the lung and liver tissues of rats exposed to incense smoke. Mol Cell Biochem. 2014 Jun;391(1-2):127-36.

44. Fa S, Larsen TV, Bilde K, Daugaard TF, Ernst EH, Lykke-Hartmann K, Olesen RH, Mamsen LS, Ernst E, Larsen A, Nielsen AL. Changes in first trimester fetal CYP1A1 and AHRR DNA methylation and mRNA expression in response to exposure to maternal cigarette smoking. Environ Toxicol Pharmacol. 2018 Jan;57:19-27.

45. Reddy RK, Charles WN, Sklavounos A, Dutt A, Seed PT, Khajuria A. The effect of smoking on COVID-19 severity: A systematic review and meta-analysis. J Med Virol. 2021 Feb;93(2):1045-1056.

46. Vardavas CI, Nikitara K. COVID-19 and smoking: A systematic review of the evidence. Tob Induc Dis. 2020 Mar 20;18:20.

47. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009 Jul 21;339:b2700.

48. Moroz VV, Smelaia TV, Salnikova LE, Golubev AM, Rubanovich AV. [Genetic study of predisposition to community-acquired pneumonia]. Vestn Ross Akad Med Nauk. 2011;(11):12-6. Russian.

49. Smelaya T V, Kuzovlev AN, Moroz VV, Golubev A M, Belopolskaya O B, Salnikova LE. Search for Common Molecular Genetic Markers of Nosocomial Pneumonia and Acute Respiratory Distress Syndrome. General Reanimatology 2015, 11, (3), 24-38.

50. Egger M, Smith GD, and Altman DG. Meta-analysis in Stata. Systematic reviews in health care: meta-analysis in context. 2001, BMJ Books (ISBN-13:978-0-7279-1488-0) (pp. 347-372.).

51. Support Stata Technical, StataCorp. 2019. Citing Stata software, documentation, and FAQs. Stata Statistical Software: 2019, Release 16.College Station, TX: StataCorp LLC.

52. Cochrane WG. The combination of estimates from different experiments. Biometrics 1954, 10, 101-129.

53. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986 Sep;7(3):177-88.

54. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 6;327(7414):557-60. doi: 10.1136/bmj.327.7414.557.

55. Light RJ, Singer JD, Willett JB. The visual presentation and interpretation of meta-analysis. In: Cooper H, Hedges LV, editors.The handbook of research synthesis. 1994, New York: Russell Sage Foundation; 1994., pp. 439-454.

56. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994 Dec;50(4):1088-101.

57. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997 Sep 13;315(7109):629-34.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

58. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol. 2001 Oct;54(10):1046-55.

59. GA Wells, B Shea, D O'Connell, J Peterson, V Welch, M Losos, P Tugwell. The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Non randomised Studies in Meta-Analyses. University of Ottawa, 2001. 2001.

60. WHO team. Estimating mortality from COVID-19. WHO Newsletter 2020, 4.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.03.21258106; this version posted June 4, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Identification

Records identified through database search (N=1406)
(PubMed, N=7; Web of Science, N= 9; Science direct, N=245; Worldwide science.org, N=1145)

Duplicate records (N=30)

Screening

Abstracts of 1376 records reviewed

Records excluded based on abstract (N = 1292)
- Review, systematic review, conference paper & meta-analysis, N=53
- Animal model or cell line studies, N=12
- Articles on infectious disease other than Pneumonia, N=1107
- Articles not related to CYP1A1, N=120

Eligibility

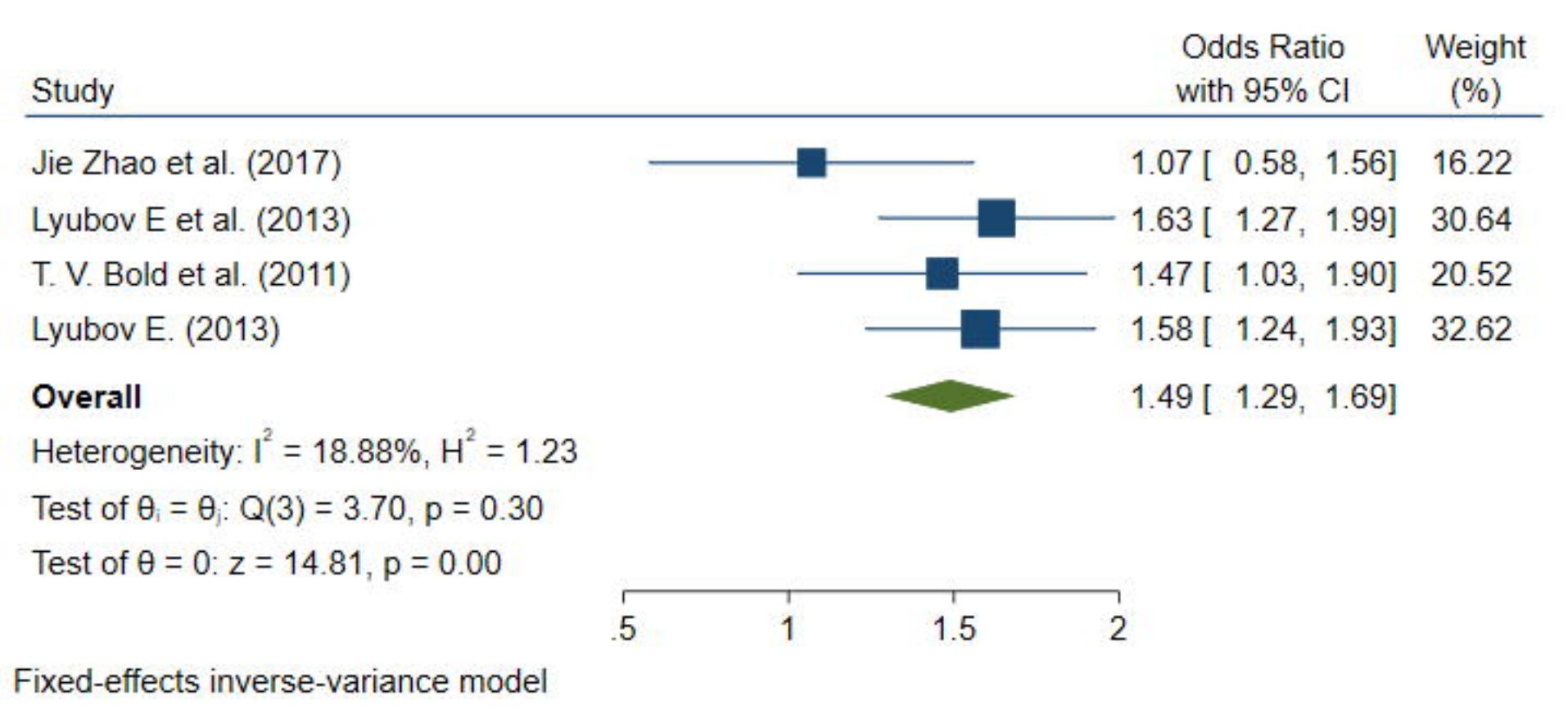
Full text of 84 records reviewed

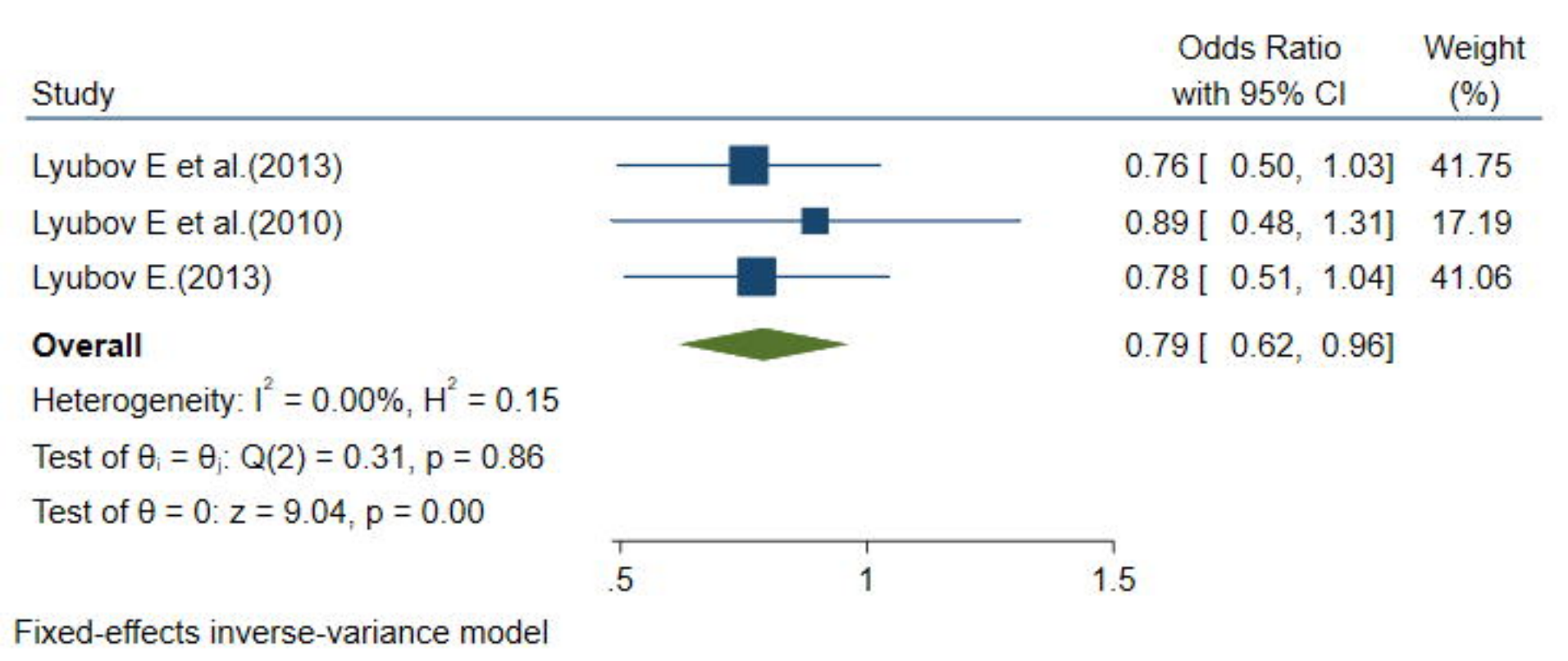
Records excluded based on full text (N = 76)
- Pneumonia genetic studies not related to CYP1A1, N=17
- Epidemiological studies, N=59

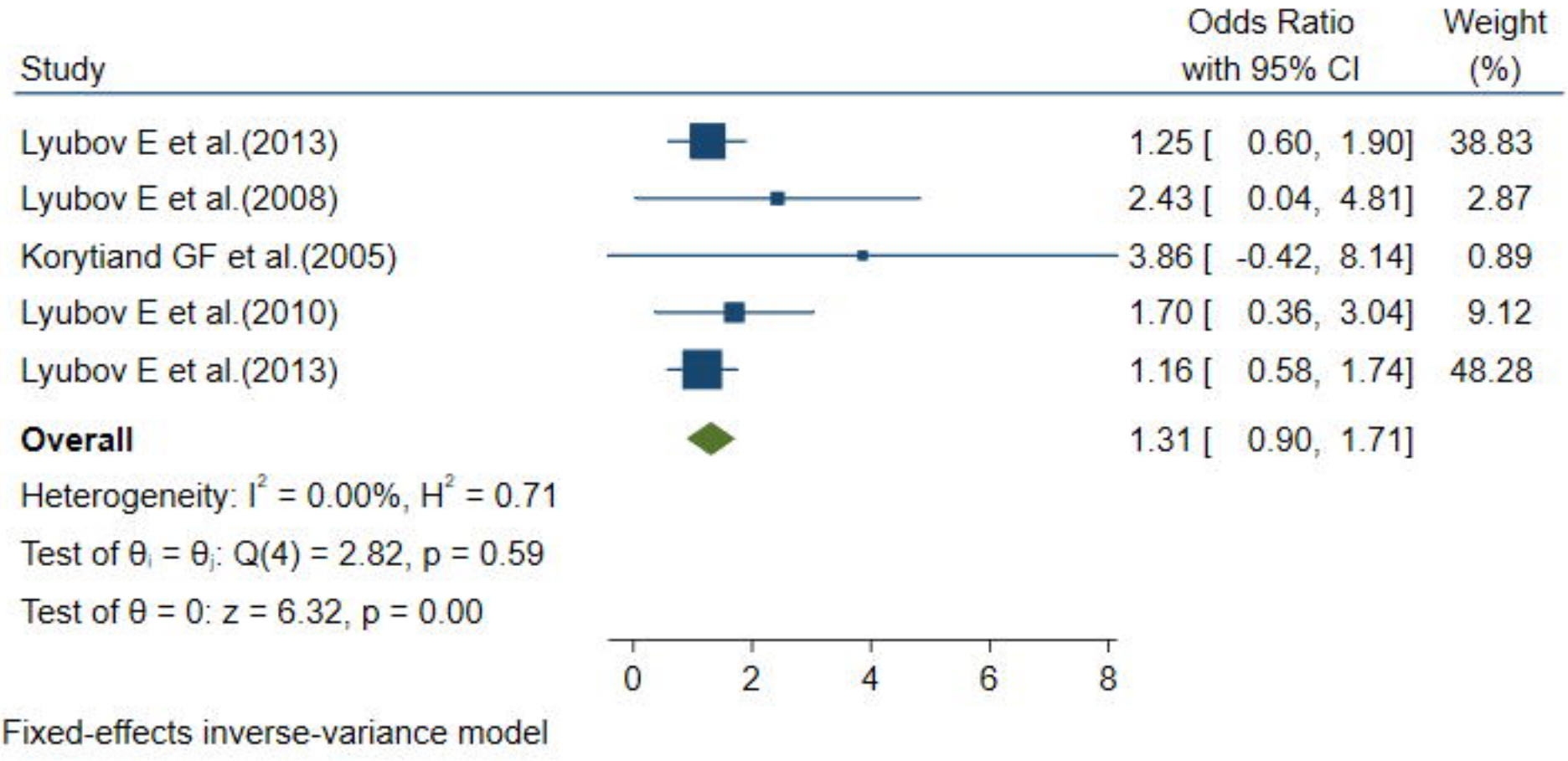
Articles included from cross-references (N = 2)

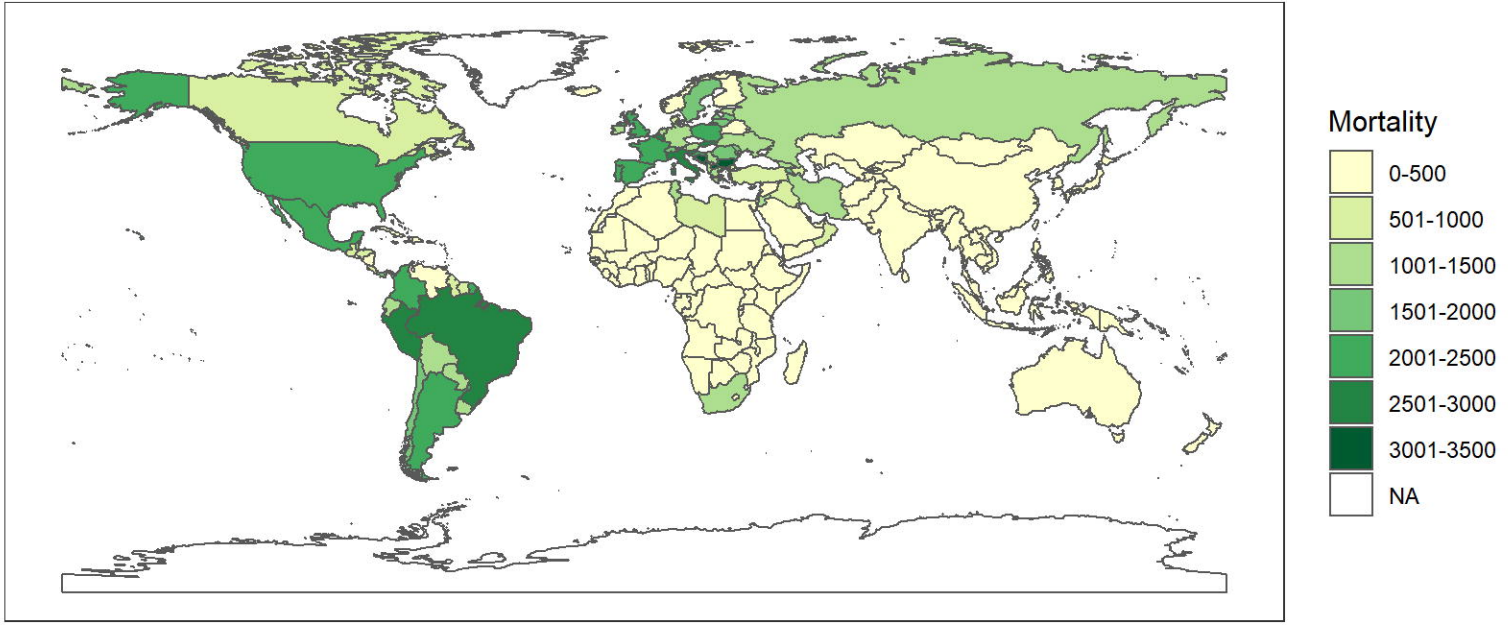
Included

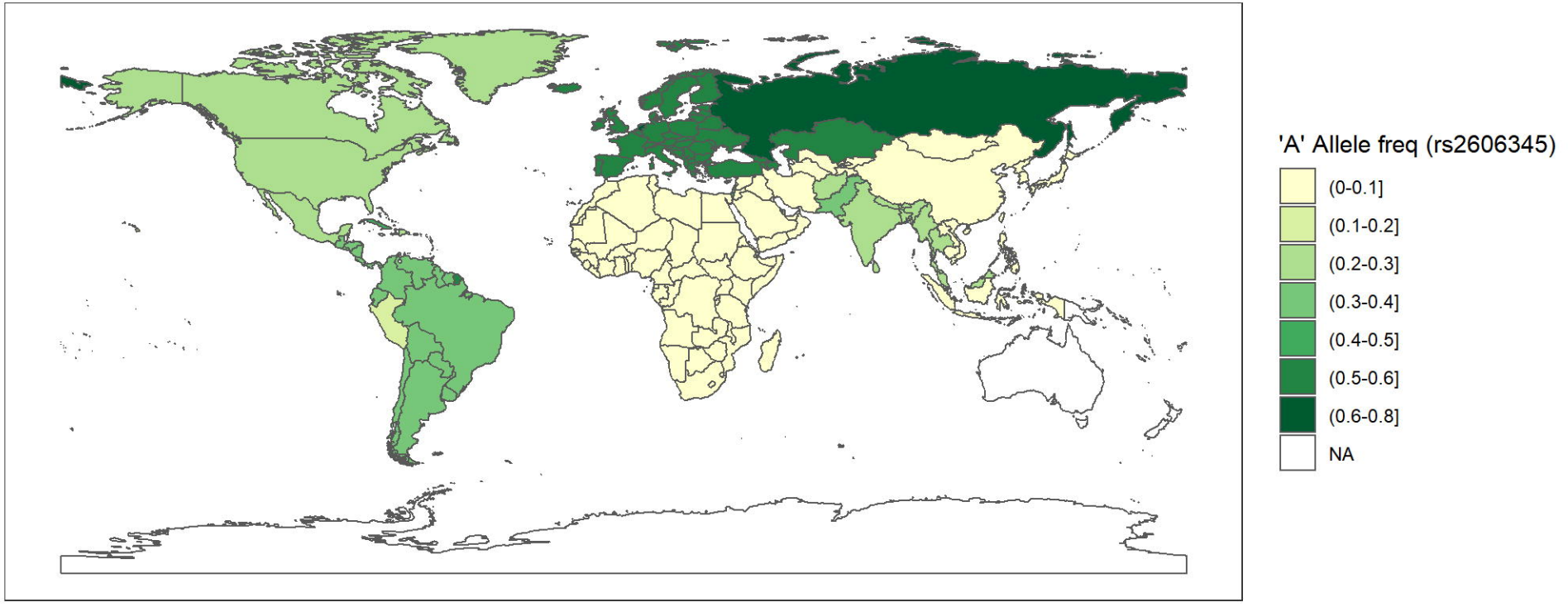
10 studies were included in meta-analysis (N = 10)

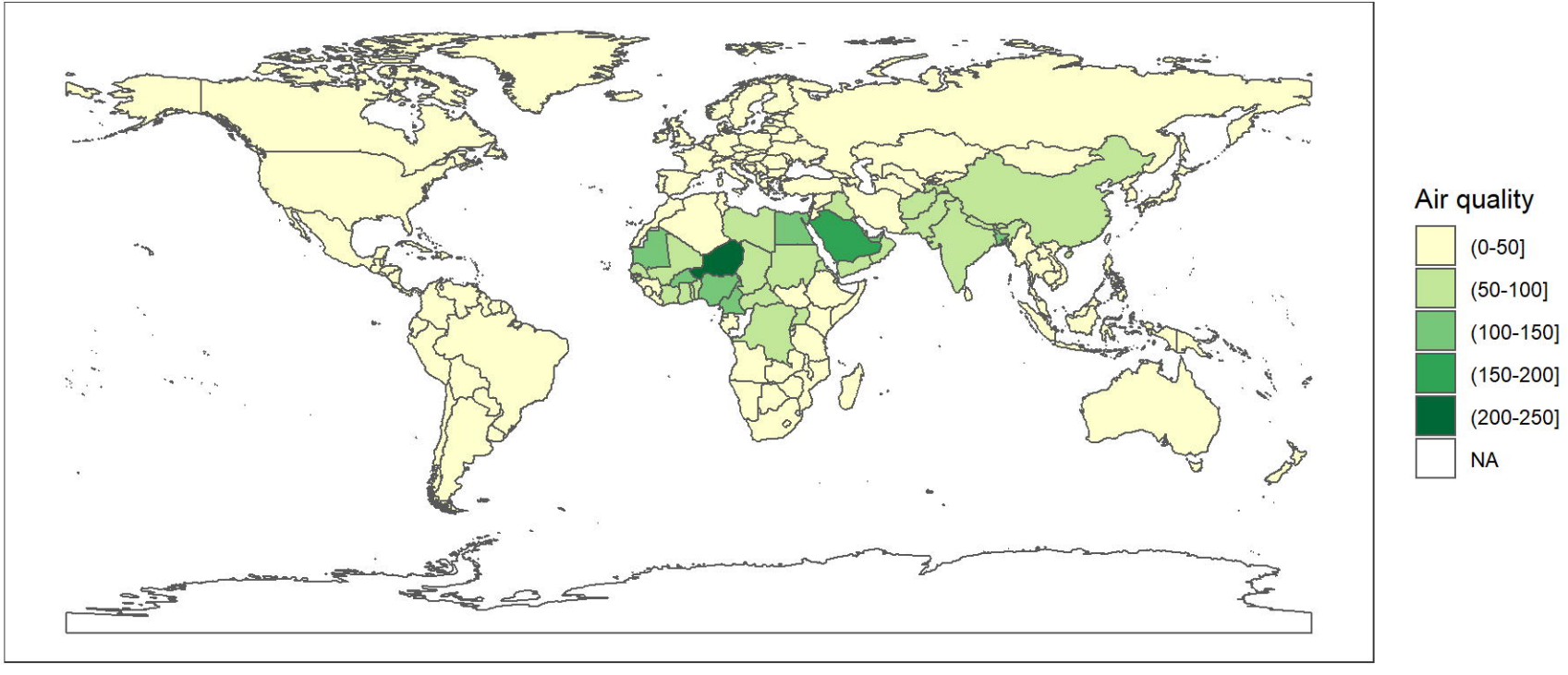

Study
Odds Ratio with 95% CI
Weight (%)
Jie Zhao et al. (2017)
1.07 [0.58, 1.56] 16.22
Lyubov E et al. (2013)
1.63 [1.27, 1.99] 30.64
T. V. Bold et al. (2011)
1.47 [1.03, 1.90] 20.52
Lyubov E. (2013)
1.58 [1.24, 1.93] 32.62
Overall
1.49 [1.29, 1.69]
Heterogeneity: I^2 = 18.88%, H^2 = 1.23
Test of $\theta_i = \theta_j$: Q(3) = 3.70, p = 0.30
Test of θ = 0: z = 14.81, p = 0.00
.5
1
1.5
2
Fixed-effects inverse-variance model

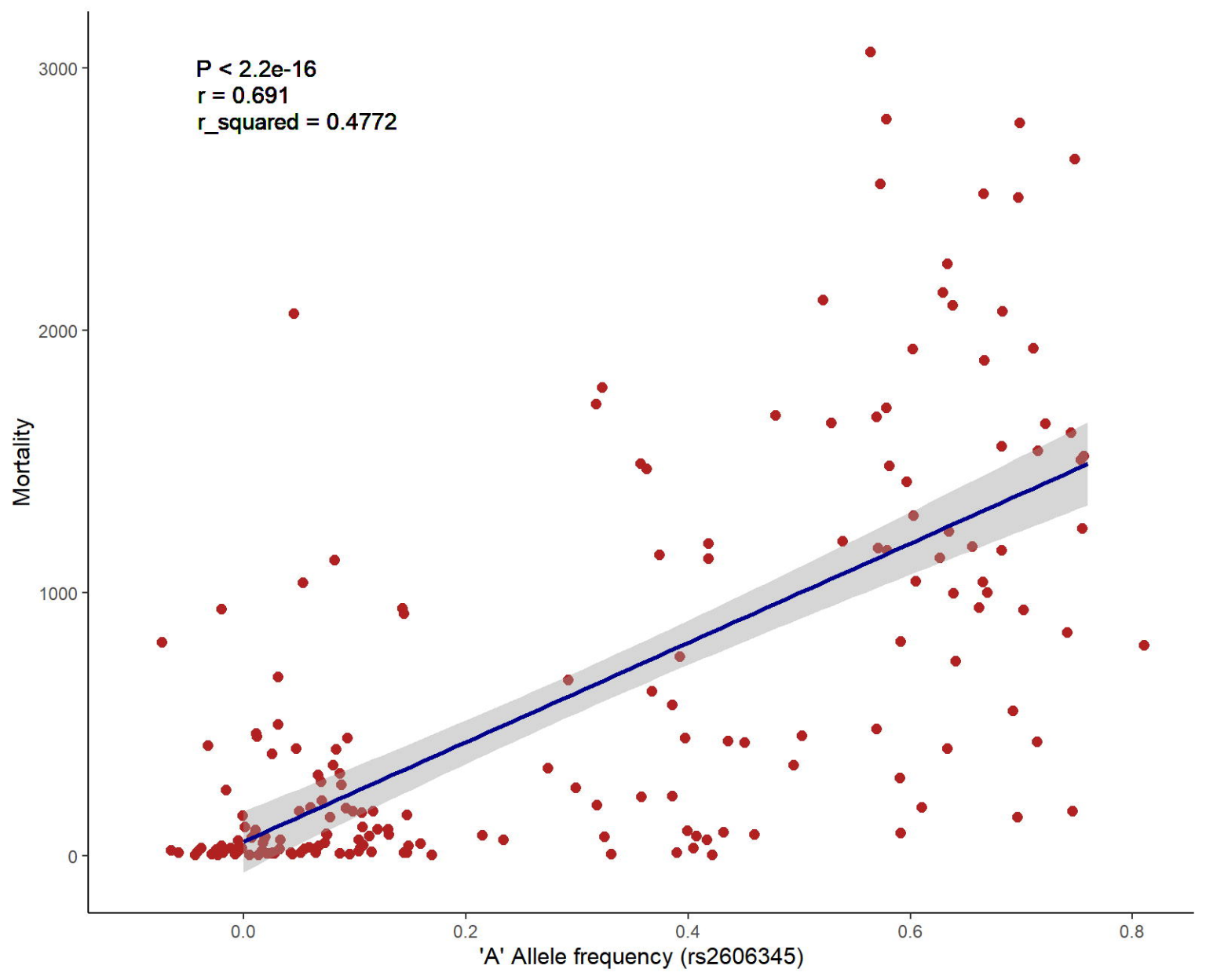

Study
Odds Ratio with 95% CI
Weight (%)
Lyubov E et al.(2013)
0.76 [0.50, 1.03] 41.75
Lyubov E et al.(2010)
0.89 [0.48, 1.31] 17.19
Lyubov E.(2013)
0.78 [0.51, 1.04] 41.06
Overall
0.79 [0.62, 0.96]
Heterogeneity: I^2 = 0.00%, H^2 = 0.15
Test of $\theta_i = \theta_j$: Q(2) = 0.31, p = 0.86
Test of θ = 0: z = 9.04, p = 0.00
.5
1
1.5
Fixed-effects inverse-variance model

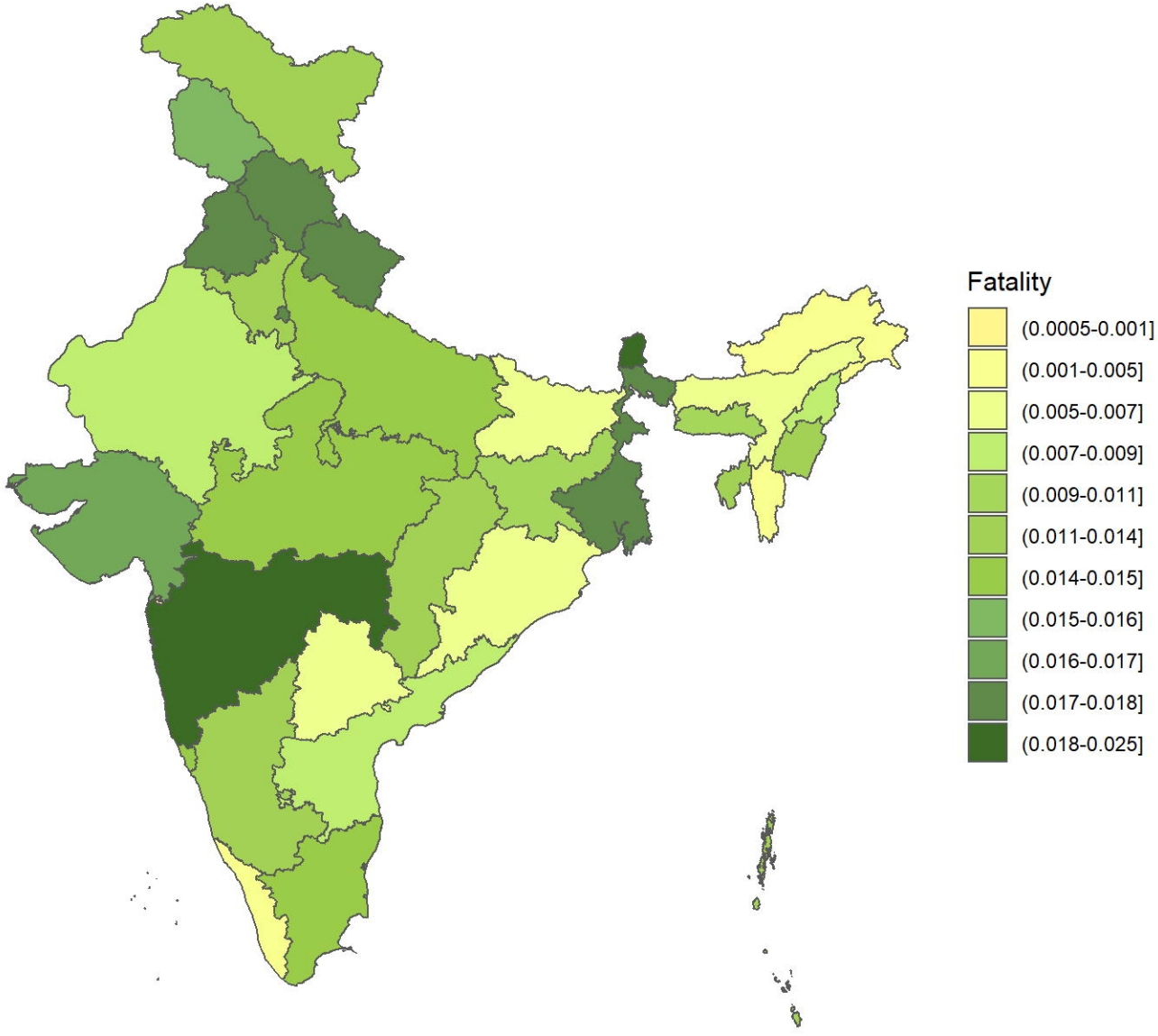

Study
Odds Ratio with 95% CI
Weight (%)
Lyubov E et al.(2013)
1.25 [0.60, 1.90] 38.83
Lyubov E et al.(2008)
2.43 [0.04, 4.81] 2.87
Korytiand GF et al.(2005)
3.86 [-0.42, 8.14] 0.89
Lyubov E et al.(2010)
1.70 [0.36, 3.04] 9.12
Lyubov E et al.(2013)
1.16 [0.58, 1.74] 48.28
Overall
1.31 [0.90, 1.71]
Heterogeneity: I^2 = 0.00%, H^2 = 0.71
Test of $\theta_i = \theta_j$: Q(4) = 2.82, p = 0.59
Test of θ = 0: z = 6.32, p = 0.00
0
2
4
6
8
Fixed-effects inverse-variance model

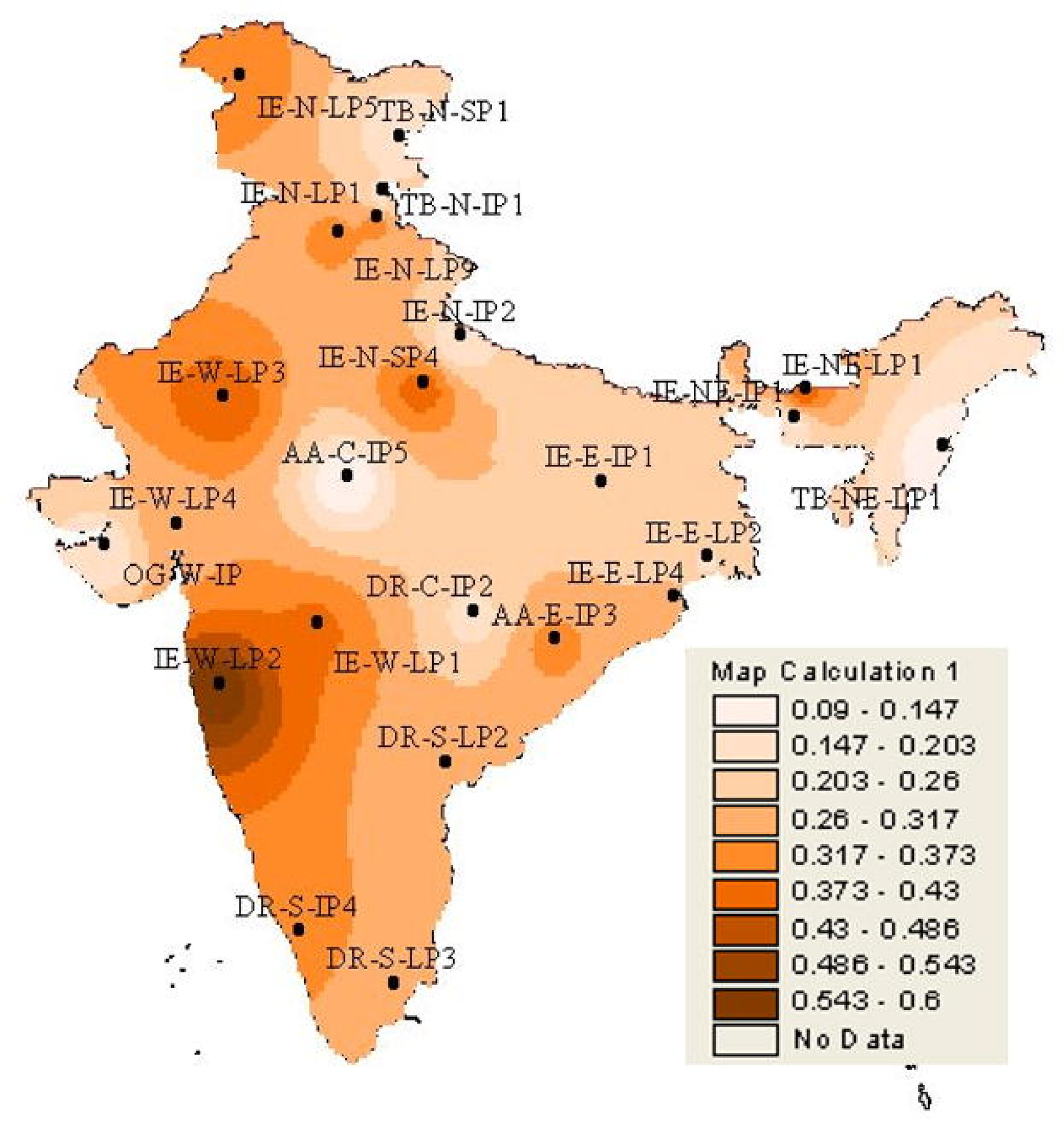

Mortality
0-500
501-1000
1001-1500
1501-2000
2001-2500
2501-3000
3001-3500
NA

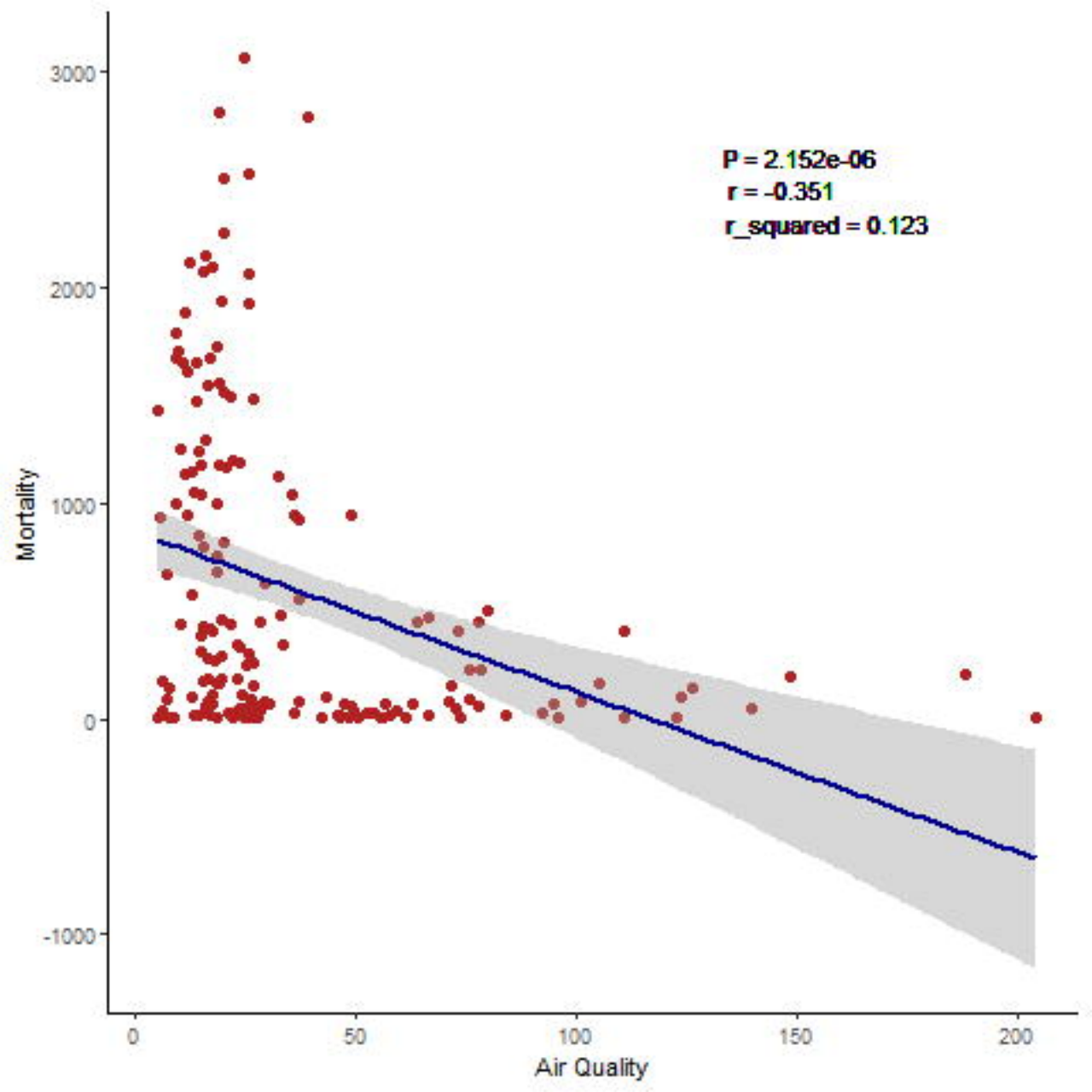

'A' Allele freq (rs2606345)
(0-0.1]
(0.1-0.2]
(0.2-0.3]
(0.3-0.4]
(0.4-0.5]
(0.5-0.6]
(0.6-0.8]
NA

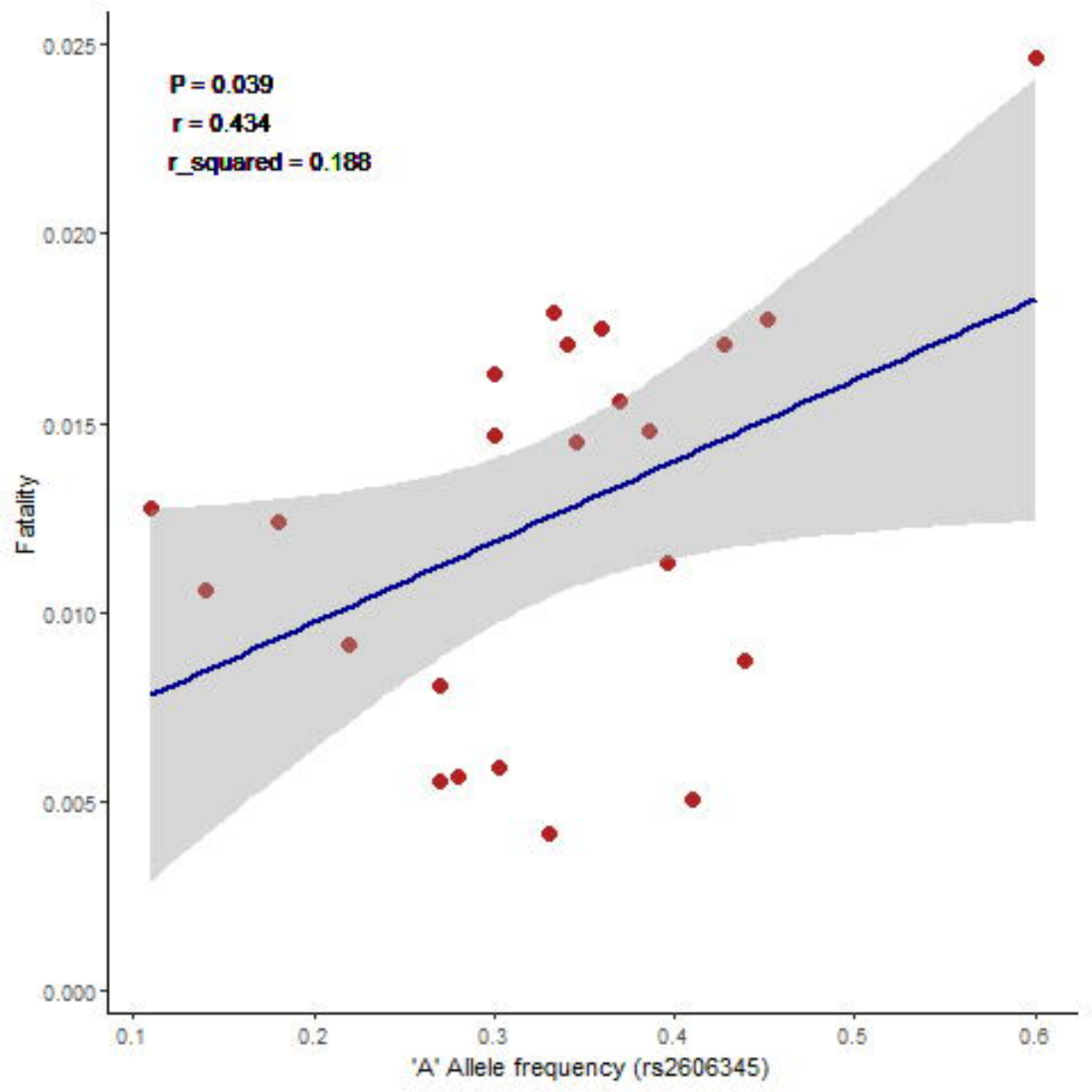

Air quality
(0-50]
(50-100]
(100-150]
(150-200]
(200-250]
NA


P < 2.2e-16
r = 0.691
r_squared = 0.4772
Mortality
3000
2000
1000
0
0.0
0.2
0.4
0.6
0.8
'A' Allele frequency (rs2606345)


Fatality
(0.0005-0.001]
(0.001-0.005]
(0.005-0.007]
(0.007-0.009]
(0.009-0.011]
(0.011-0.014]
(0.014-0.015]
(0.015-0.016]
(0.016-0.017]
(0.017-0.018]
(0.018-0.025]


IE-N-LP5
TB-N-SP1
IE-N-LP1
TB-N-IP1
IE-N-LP9
IE-N-IP2
IE-W-LP3
IE-N-SP4
IE-NE-LP1
IE-NE-IP1
AA-C-IP5
IE-E-IP1
IE-W-LP4
TB-NE-IP1
IE-E-LP2
OG-W-IP
IE-E-LP4
DR-C-IP2
AA-E-IP3
IE-W-LP2
IE-W-LP1
DR-S-LP2
DR-S-IP4
DR-S-LP3
Map Calculation 1
0.09 - 0.147
0.147 - 0.203
0.203 - 0.26
0.26 - 0.317
0.317 - 0.373
0.373 - 0.43
0.43 - 0.486
0.486 - 0.543
0.543 - 0.6
No Data


P = 2.152e-06
r = -0.351
r_squared = 0.123
Mortality
3000
2000
1000
0
-1000
0
50
100
150
200
Air Quality


P = 0.039
r = 0.434
r_squared = 0.188
Fatality
0.025
0.020
0.015
0.010
0.005
0.000
0.1
0.2
0.3
0.4
0.5
0.6
'A' Allele frequency (rs2606345)